Effects of voice training on phonetograms and maximum phonation times in female speech therapy students

Arend M. Sulter and Jorien M. Meijer

Department of Otorhinolaryngology, University Hospital Groningen, Groningen, the Netherlands

Submitted

INTRODUCTION

Voice training affects morphological aspects of, as well as the control over the voice source. With specific instructions, benign lesions of the vocal folds, specifically vocal fold nodules (1-3), decrease in size (4-6), and in some cases the closure of vocal folds can be improved (7, 8). The instructions given during voice therapy or training modify and improve laryngeal muscle strength, tone, balance, and stamina (9). By establishing conditions for a healthy vocal fold cover, these instructions result in ameliorating the symptoms of benign lesions and in many cases prevent a recurrence (10-12). As trained subjects, singers also elicit effects of the improved control over the voice source. A number of studies exemplify the differences between singers and nonsingers with respect to motor control (see Murry & Caligiuri) (13).

Improved laryngeal muscle strength and tone, as well as improved balance among laryngeal muscle effort, respiratory effort and control, might result in an increase in vocal capacities. These capacities can be visualized in a phonetogram (14). Figure 1 gives an example of a phonetogram with lines representing frequency and intensity ranges. Comparing phonetograms of
Figure 1. An example of a "normal" male phonetogram. Along the x-axis the frequency scale is plotted (32.7 - 2096 Hz) and the intensity level is given along the y-axis (40 - 120 dB). Note the dip in the loud phonation contour at about 400 Hz. This local minimum exhibits the transition of chest register to falsetto register. The lines in the vertical and horizontal direction indicate the dynamic and frequency range, respectively.

Speech training only increased intensities are to be expected, as phonation with fundamental frequencies beyond pitches used in normal speech is not practiced.

Another method to analyze voice function is measuring maximum phonation times (19). The s/z ratio has previously been employed as an indicator of laryngeal pathology (20). As the maximum phonation time of the /s/ reflects the expiratory flow control and the /z/ indicates glottal resistance, the s/z ratio might change after voice training.

Most previous investigations of voice training effects are cross-sectional studies. A possible a priori difference between groups, independent of voice training, therefore can not be excluded. A prospective study is more suitable for analyzing effects of voice training.

The specific content of voice training depends on the intended goal. Training given to voice patients is directed toward curing or ameliorating pathology, correcting pathogenic voicing patterns, as well as optimizing the use of voice capacities. The optimization is achieved by correcting breath control, promoting relaxed phonation and employing resonance potencies by adjusting vocal tract anatomy.

Singing differs from speech with respect to intensity and frequency ranges, breath management and neuromyogenic control (21). Voice training given to singers focuses on improving these aspects of phonation and therefore differs from speech training.
In this study effects of specific speech training given to female speech therapy students are analyzed by comparing phonetograms and maximum phonation times before (PRE) and after (POST) two and a half years of training.

METHODS

Subjects
A group of 25 female students (age 17 - 21 years, mean age 18.4 years, standard deviation (SD) 1.20 years) was investigated before starting the study of speech therapy at the Training School for Speech Therapist, Groningen. After two and one half years of education the same investigation, among others consisting of phonetography and measuring phonation times, was repeated. Pathology of vocal folds was excluded by laryngostroboscopic examination of the students. The effect of ageing of the students on the measured variables was considered by analyzing a data base with normative data on a comparable age cohort (17).

Voice training
Voice training entailed practicing several methods during speech therapy study in Groningen. The goal of voice training was to attain a clear voice carriage with adequate (normal tonus) bodily posture and diaphragmatic breathing. The training methods are described in short beneath. For a more detailed description of the exercises the reader is referred to the literature (see references).

The exercises of Coblenzer (22, 23) required coupling breathing, phonation and rhythmic movements. Essential in this method is the bodily posture (eutonus), comprising the balance among laryngeal muscle effort, respiratory effort and control, and supraglottal modification of the laryngeal tone. This posture enables maximum vocal performances with a minimum of effort. When the right balance in tension has been found, the attention is directed toward breath management. The last part aims at learning to use reflexory movements of the diaphragm to provide automatic inhalation.

The resonance method (24, 25) required a proper balance between stress and relaxation during phonation. Manipulating resonance characteristics by adjusting vocal tract configuration in combination with a low pharyngeal position, the voice gains intensity with clear carriage.

The Smith Accent Method (26) aims at the development of abdominal-
diaphragmatic breathing using accentuated rhythmic movements. This facilitates optimal breath management, resulting in a firm glottal closure. The gist of the method is practicing in chest register with vocal folds that are short and lax, and therefore have a large vibrating mass.

The nasalizing method (27) derives its name from a striking and characteristic part of the method: the nasalization of sound. The nasal sound originates from a completely relaxed larynx and vocal tract, where the air flows both through mouth and nose. The soft palate hangs downward loosely, which results in a relaxed "lifting" mechanism. The relaxation has a beneficial effect upon voice function. The resonant space is larger because of a wider and longer vocal tract, and the resonant quality is improved by the relative absence of tension in the larynx and vocal tract. The possibilities for modulating articulations are improved in combination with less eminent vocal and speech fatigue.

The students were all equally trained during a period of two and a half years. The forementioned methods were exercised and practiced for a total of 260 hours of education per student.

Phonetography

Equipment. Phonetograms were registered in a sound treated room with "living room acoustics" (28), using an FST-II phonetometer. This phonetometer and its operation are described in detail in Sulter et al. (17).

All phnetograms were registered by two investigators, both familiar with the equipment and phonetographic procedure, and having had more than five years musical training and experience. Both investigators used a standard set of detailed instructions for the students.

Phonographic Procedure. Phonotography recommendations by the UEP were followed (28). The direction and distance (30 cm) of the student's mouth to the microphone was carefully controlled during the procedure. The students were tested using the vowel /a/. Phonetogram points were collected, starting the acquisition at the mean speaking fundamental frequency, followed by the low frequencies and ending with high frequencies. The mean speaking fundamental frequency was determined by asking the students to count from one to ten. Then the investigator's imitation of that frequency was measured with the phonetometer. A reproducible phonation at a given frequency with a minimum phonation time of one second, to insure a stabilized sound intensity production and correct measurement, was required for accepting a phnetogram point. During actual registration, the student, if necessary, was at first guided in matching the target frequency, and after this the minimum and maximum intensity were registered. The students were instructed to produce phonations at the physiologic boundaries, without, of course, injuring the voice during
Effects of Voice Training

maximal intensity. The frequency range was sampled at four frequencies per octave, basically at the tones c-e-g-a, e.g., C4-E4-G4-A4 (in this octave, 262, 330, 392 and 440 Hz). At the upper and lower ends of the range, shorter frequency intervals of semitones were chosen. The phonetograms were acquired within a ten to twenty minute time period.

Phonetogram Analysis. For reasons of comparison and establishing differences between phonatory capacities a standardized approach was used. Phonetograms were analyzed using two different methods, which are described in detail in Sulter et al. (17, 29). What follows is a brief description of these methods.

Rescaling method. The rescaling method determines dynamic ranges at fixed relative distances along a student's (individually differing) frequency range (see Figure 1). Retrieving phonetogram points from individual data files, each frequency range was rescaled to 100% with a specially developed computer program. At 10% intervals the minimum and maximum intensity was calculated by interpolation, yielding eleven values for both minimum and maximum sound pressure levels. By averaging the SPL values at each interval for a number of students, normative data on the dynamic ranges were established.

Conjoint frequency and intensity analysis of the phonetogram. Another approach for the analysis of phonetograms concentrates on the phonetogram features shape, area and Weighted Dynamic Range and Central Position. The main difference between this approach and the rescaling method is that it derives variables without distortion of the shape.

The shape of a phonetogram is described with so-called Fourier Descriptors (FDs), thus enabling quantification (29, 30). Changes in shape can be notified with these FDs. Shape can also be described by contour regularity, which is the quotient of enclosed area and squared perimeter, producing a dimensionless value. The slope of a phonetogram is determined by drawing a line with minimal distance to all points through the phonetogram and calculating the angle between this line and x-axis.

The enclosed area gives the quotient of the voice field area and the rectangle with coordinates 40 and 110 dB, and 32.7 and 2096 Hz. The frequency range is derived by subtracting the lowest phonated fundamental frequency from the highest frequency.

A Weighted Dynamic Range (WDR) is determined in the modal register at four frequencies which are related to the mean speaking fundamental

¹The four sampled frequencies in the modal register are: mean speaking fundamental frequency (mff) minus three semitones; mff; mff plus half an octave; and mff plus an octave. In this study the mff was standardized at 220
frequency. Because sound intensities around 75 dB are relatively more important in normal speech than very soft and loud intensities, a logarithmic weighting factor was used, giving greater weight to the intensities in accordance to how close they are to 75 dB. Besides the WDR, the Central Position (CP) of this range is calculated. For a detailed description the reader is referred to Sulter et al. (29).

Both these methods for analyzing phonetograms were implemented in a computer program written in the ASYST language (ASYST, Macmillan software company) and running in DOS environment.

Maximum phonation times

After a demonstration by the examiner, the students were instructed to inhale and to produce the vocalized consonant /z/ for as long as possible at a comfortable pitch and loudness level. The same instructions were given for the production of a sustained /s/. Each student produced the consonants twice. Productions were measured with a stopwatch with an accuracy of 0.1 s. The longest phonation of each consonant was used for further analyses.

Statistical analysis

To analyze the effect of voice training, comparisons were made between PRE and POST phonetograms, as well as between maximum phonation times. Differences in variables resulting from the rescaling method were analyzed with two-way analysis of variance (ANOVA) with rescaled frequency value and voice training (PRE/POST) as factors. To detect significant differences among factor levels, post hoc Tukey HSD tests were performed. Differences between variables resulting from the phonetographical conjoint analysis method, as well as differences in maximum phonation times were analyzed with paired t-tests. Linear regression was used to analyze the influence of age on phonetographical variables and maximum phonation times. A probability level of 0.05 was used to reject the null hypothesis, which posits no difference among variables under investigation.

RESULTS

First the results of comparing phonetograms will be given, followed by a presentation of the results of maximum phonation times.

Phonetograms

Rescaling Method. After averaging the 25 phonetograms with the rescaling
method, descriptive statistics can be derived for each rescaled frequency value and vocal intensity. These average rescaled frequency values will first be discussed, and then the average minimum and maximum intensity levels will be presented. Inferential statistics are used to establish differences in untrained and trained phonetograms.

Rescaled frequency values. Table 1 gives average frequencies with standard deviations at the 10% rescaled frequency values.

An average frequency range from 158.4 Hz to 1255.7 Hz was measured in the untrained female students (PRE). After two and a half years of education (POST) average frequency values ranging from 157.9 to 1212.5 were established. Thus the trained students have a reduced frequency range and phonate slightly lower at the 100% rescaled frequency value; however, this difference is less than one semitone.

The standard deviation shows an absolute increase in frequency as the rescaled value rises. In semitones this amounts to about 2 semitones at the 0% rescaled frequency values and 3 semitones at the 100% rescaled frequency values.

Sound intensity levels. Table 2 gives the average maximum and minimum phonation intensities, together with standard deviations at the 10% rescaled frequency values. Using the average frequencies of Table 1, Figure 2 was constructed, visualizing the average phonetograms for untrained (PRE) and trained (POST) students.

The standard deviations for frequencies and intensities are represented as whiskers in both plots. Apart from the values near the extreme rescaled frequency values, the standard deviations of intensities are generally less than 6 dB. With increasing frequency a rising intensity level is present for the soft phonation contour in both the untrained and trained average

<table>
<thead>
<tr>
<th>Frequency level</th>
<th>PRE</th>
<th>POST</th>
</tr>
</thead>
<tbody>
<tr>
<td>0% mean (Hz)</td>
<td>158.4</td>
<td>157.9</td>
</tr>
<tr>
<td>SD (Hz)</td>
<td>23.78</td>
<td>19.01</td>
</tr>
<tr>
<td>10% mean (Hz)</td>
<td>194.2</td>
<td>193.2</td>
</tr>
<tr>
<td>SD (Hz)</td>
<td>25.77</td>
<td>20.83</td>
</tr>
<tr>
<td>20% mean (Hz)</td>
<td>238.2</td>
<td>236.4</td>
</tr>
<tr>
<td>SD (Hz)</td>
<td>28.49</td>
<td>23.82</td>
</tr>
<tr>
<td>30% mean (Hz)</td>
<td>292.6</td>
<td>289.4</td>
</tr>
<tr>
<td>SD (Hz)</td>
<td>32.69</td>
<td>29.01</td>
</tr>
<tr>
<td>40% mean (Hz)</td>
<td>359.5</td>
<td>354.4</td>
</tr>
<tr>
<td>SD (Hz)</td>
<td>39.85</td>
<td>37.66</td>
</tr>
<tr>
<td>50% mean (Hz)</td>
<td>442.0</td>
<td>434.3</td>
</tr>
<tr>
<td>SD (Hz)</td>
<td>51.76</td>
<td>51.14</td>
</tr>
<tr>
<td>60% mean (Hz)</td>
<td>543.7</td>
<td>532.8</td>
</tr>
<tr>
<td>SD (Hz)</td>
<td>70.26</td>
<td>70.59</td>
</tr>
<tr>
<td>70% mean (Hz)</td>
<td>669.6</td>
<td>653.6</td>
</tr>
<tr>
<td>SD (Hz)</td>
<td>98.08</td>
<td>98.31</td>
</tr>
<tr>
<td>80% mean (Hz)</td>
<td>825.1</td>
<td>802.7</td>
</tr>
<tr>
<td>SD (Hz)</td>
<td>137.58</td>
<td>136.38</td>
</tr>
<tr>
<td>90% mean (Hz)</td>
<td>1017.5</td>
<td>986.2</td>
</tr>
<tr>
<td>SD (Hz)</td>
<td>192.96</td>
<td>188.13</td>
</tr>
<tr>
<td>100% mean (Hz)</td>
<td>1255.7</td>
<td>1212.5</td>
</tr>
<tr>
<td>SD (Hz)</td>
<td>269.10</td>
<td>257.12</td>
</tr>
</tbody>
</table>

Table 1. Mean frequencies and standard deviations (SD) for consecutive frequency levels of a group of speech therapy students at the beginning of the study (PRE) and after two and a half years of education (POST).
Table 2. Maximum and minimum intensities in a group of speech therapy students at the beginning of the study (PRE) and after two and a half years of education (POST). Mean intensity levels and standard deviations (SD) are given.

<table>
<thead>
<tr>
<th>Frequency level</th>
<th>Maximum</th>
<th>Minimum</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PRE</td>
<td>POST</td>
<td>PRE</td>
</tr>
<tr>
<td>0% mean (dB)</td>
<td>66.1</td>
<td>69.4</td>
<td>56.4</td>
</tr>
<tr>
<td></td>
<td>10.80</td>
<td>9.10</td>
<td>7.38</td>
</tr>
<tr>
<td>10% mean (dB)</td>
<td>86.2</td>
<td>87.6</td>
<td>54.8</td>
</tr>
<tr>
<td></td>
<td>4.91</td>
<td>3.33</td>
<td>6.40</td>
</tr>
<tr>
<td>20% mean (dB)</td>
<td>91.8</td>
<td>95.4</td>
<td>57.3</td>
</tr>
<tr>
<td></td>
<td>3.86</td>
<td>3.55</td>
<td>5.77</td>
</tr>
<tr>
<td>30% mean (dB)</td>
<td>96.5</td>
<td>98.7</td>
<td>59.3</td>
</tr>
<tr>
<td></td>
<td>5.08</td>
<td>2.54</td>
<td>5.78</td>
</tr>
<tr>
<td>40% mean (dB)</td>
<td>99.2</td>
<td>101.1</td>
<td>61.2</td>
</tr>
<tr>
<td></td>
<td>5.47</td>
<td>4.73</td>
<td>6.10</td>
</tr>
<tr>
<td>50% mean (dB)</td>
<td>99.3</td>
<td>103.3</td>
<td>64.4</td>
</tr>
<tr>
<td></td>
<td>6.23</td>
<td>4.92</td>
<td>7.33</td>
</tr>
<tr>
<td>60% mean (dB)</td>
<td>99.1</td>
<td>104.1</td>
<td>68.7</td>
</tr>
<tr>
<td></td>
<td>5.01</td>
<td>5.44</td>
<td>7.60</td>
</tr>
<tr>
<td>70% mean (dB)</td>
<td>101.2</td>
<td>103.3</td>
<td>74.0</td>
</tr>
<tr>
<td></td>
<td>5.14</td>
<td>3.12</td>
<td>7.14</td>
</tr>
<tr>
<td>80% mean (dB)</td>
<td>104.1</td>
<td>104.4</td>
<td>80.1</td>
</tr>
<tr>
<td></td>
<td>6.05</td>
<td>4.05</td>
<td>7.84</td>
</tr>
<tr>
<td>90% mean (dB)</td>
<td>104.9</td>
<td>106.3</td>
<td>87.0</td>
</tr>
<tr>
<td></td>
<td>7.15</td>
<td>3.43</td>
<td>8.76</td>
</tr>
</tbody>
</table>

The average trained (POST) phonetogram shows better possibilities for producing soft intensities, compared to the average untrained phonetogram. A distinct difference in the loud phonation contour between the average phonetograms is the location of a relative minimum in an otherwise continuously rising loud phonation contour, which is present at the 60% rescaled frequency value (about 500 Hz) of the average untrained (PRE) phonetogram, whereas it is located at the 70% rescaled frequency value (about 600 Hz) in the average trained (POST) phonetogram. This relative minimum marks the transition from modal to falsett o register, which therefore seems to occur at a higher frequency in the trained group.

Table 3 gives the results of the ANOVA procedure applied to frequency and intensities. No significant interaction was found between the factors rescaled frequency value and training. Not surprisingly, the effect of the factor rescaled frequency value on frequency is highly significant. The factor vocal training (PRE/POST) has no significant effect on calculated frequencies. Both the factor rescaled frequency value and vocal training have significant effects on maximum and minimum intensity.

Post hoc Tukey-HSD tests with a significance level \( p=0.05 \) performed for the factor vocal training showed a significant difference in maximum intensity.
Figure 2. Averaging frequency and intensity data, mean phonetograms are plotted for a group of 25 female voice therapy students (above) at the beginning of the study and (below) after two and a half years of education. The whiskers indicate 1 standard deviation.

Table 3. Analysis of variance summary table with effects of factors frequency level and voice training on produced frequencies, and maximum and minimum intensities. No significant interaction was present between the factors.

<table>
<thead>
<tr>
<th></th>
<th>Frequency level</th>
<th>Voice training</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>p</td>
</tr>
<tr>
<td>Frequency</td>
<td>1162.9</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Maximum intensity</td>
<td>177.0</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Minimum intensity</td>
<td>211.0</td>
<td>&lt;0.001*</td>
</tr>
</tbody>
</table>

Note. df=10,539 for frequency level. df=1,548 for voice training.
* p<0.05

Conjoint Frequency and Intensity Analysis. The results obtained with the conjoint analysis method are given in Table 4. Paired t-tests were performed to detect significant differences between average variables from untrained and trained phonetograms.

Shape. Before and after the period of vocal training the average shape shows Fourier Descriptors with the previously described pattern of alternating low and high values (17). For the higher order FDs the high values diminish and level off (see Table 4). A significant difference in average shape was established for the first (FD₁), second (FD₂), as well as the seventh
Fourier Descriptor (FD). The contour regularity increased significantly from 0.20 to 0.24 after voice training. The slope of the phonetogram did not change significantly.
Table 4. Analyzed characteristics of phonetograms of a group of speech therapy students at the beginning of the study (PRE) and after two and a half years of education (POST). Mean values and standard deviations (SD) are given. FD = Fourier Descriptor, WDR = Weighted Dynamic Range, CP = Central Position, mff = mean speaking fundamental frequency, st = semitones, oct = octaves. Note. df = 24. * p<0.05.
Area. The enclosed area shows a highly significant increase from 0.20 to 0.27. In contrast, the frequency range did not change significantly.

WDR and CP. On all four frequency positions the WDR increases highly significantly with vocal training (see Table 4). This difference is most significant at the mean speaking fundamental frequency plus half an octave (WDR_{mff+1/2oct}).

Before and after vocal training, the calculated CP of the WDR show both negative and positive values, implying that the centre of the intensity range is below or above the reference intensity of 75 dB, respectively. The transition (with increasing frequency within the phonetogram) from a negative to a positive value occurs between the sampled frequencies mff and mff+1/2oct. The CP of the average trained phonetogram is located at a lower level at all four frequency intervals. The difference in location of the CP is significant at the sampled frequencies CP_{mff+1/2oct} and CP_{mff+1octave}.

**Maximum Phonation Times**

Table 5 summarizes the results of produced maximum phonation times. Surprisingly, phonation time of the vocalized consonant /z/ decreased with vocal training. The average maximum phonation time of the voiceless consonant /s/, however, increased and showed the largest value. The standard deviation for the phonation time of /s/ is large in comparison with that of /z/ and expresses the difference across students in expiratory control. Paired t-tests revealed a significant decrease in phonation time for /z/ and a significant increase for the s/z ratio from 1.11 to 1.45.

**DISCUSSION**

Effects of voice training are discussed with respect to phonetograms, followed by a discussion of presented maximum phonation times.
Phonetograms

Clear differences were established between average phonetograms before and after vocal training. As demonstrated by the variables range, enclosed area, and Weighted Dynamic Range, voice capacities are increased with respect to dynamic ranges (see Tables 2 and 4). This increment offers enhanced possibilities to modulate intensity during speech. The lower Central Positions (see Table 4) indicate that the minimum intensities can be produced softer. The minimum intensity profile has a direct relation with threshold pressure (31, 32). Threshold pressure is, among others, determined by fundamental frequency, vocal fold thickness and longitudinal tension. A successful implementation of the Smith accent method should lead to phonation with vocal folds with increased thickness, which thus lowers threshold pressure and thereby increases soft voice capacities. Another condition favorable for producing soft phonations is increased breath support and expiratory control. This improved control should follow from the exercises of Coblenzer. A softer minimum intensity profile can therefore be regarded as a specific result of voice training.

The maximum intensity profile shows a significant increase in loudness. The sound pressure level measured outside the mouth is the product of both laryngeal source and supralaryngeal resonance characteristics. The maximum loudness at the source level is determined by the maximum subglottal pressure a subject is able -- or willing -- to produce. This maximum is limited by various factors, but most importantly is phonatory instability pressure (31). Above this pressure the voice quality deteriorates. Subjects, therefore, avoid passing this threshold pressure. Although subglottal pressure is the main component determining SPL, another important laryngeal condition promoting loud phonations is "flow phonation". Flow phonation is achieved in a relaxed mode of phonation (33). A successful application of the Smith accent method shows a positive relation between increased airflow and SPL (34). Therefore, trained subjects should be able to produce louder phonations. The trained subjects might also benefit from an increased awareness of their own voice capacities. This awareness diminishes the fear of damaging vocal structures by phonation at high sound intensities, which stimulates a deliberate exploration of loud voice capacities.

The vocal tract supplies optimal acoustical characteristics for the power transfer from voice source to mouth opening. The resonances in the vocal tract give an important contribution to the overall SPL. Various parts of the methods involved in voice training, such as the resonance method and the nasalizing method, are aimed at improving the impedance match between the glottis and free space (31).

The observed differences in dynamic range might also be attributed to the time period of two and a half years involved in this study and a possible
influence of ageing of the students. Regression analysis in a data base with normative data was used to examine the effect of age on maximum and minimum intensity. In a cohort of 83 untrained female subjects (age between 17 and 25, mean age 18.9, SD = 1.61 years) a significant negative effect (p=0.0038) of age was established on maximum intensity, indicating a decrease in maximum intensity with age. A significant negative effect of age (p=0.0161) with a regression coefficient of -0.79 was found for the minimum intensity. Thus, in the investigated group of students a decrease in minimum intensity of 2.0 dB might be expected, which is much less than the measured 7.2 dB. The increase in loud and soft voice capacities can, therefore, not be explained by ageing of the students.

Pertinent studies focusing on phonetographical differences between trained and untrained subjects showed, to a varying extent, increased dynamic ranges in the trained ones. Åkerlund et al. (15) established a significantly increased loud phonation contour, while no difference was found in the soft phonation contour. In contrast, Sulter et al. (17) found increased soft phonation capacities in trained amateur singers, while there was no difference in loud phonation contour. Awan (16) established a significant increase in both loud and soft capacities. These studies also established an increase in frequency range in the trained subjects. In the present study such a difference was not found. The unchanged frequency range might be due to the fact that the training methods are not aimed at improving singing capacities, which would have implied training phonation in falsetto, a register normally not used in Dutch speech. The PRE and POST frequency ranges are in agreement with the frequency range of a large group of untrained female subjects (17).

Besides the minute analysis of the dynamic range in the speaking voice area, the phonetographical conjoint analysis method also enables a numerical evaluation of shapes of phonetograms. Studying Figure 2, the average trained phonetogram clearly shows the larger enclosed area. In both average phonetograms there is a local depression, related to a register shift, in an otherwise rising loud phonation contour; however, the location along the frequency range is shifted upward in frequency in the trained phonetogram. Three of the twelve FDs used in this study, namely FD₁, FD₂, and FD₇, show a significant difference between the PRE and POST phonetogram. The lower value for FD₁ after vocal training represents the more circular shape of the average phonetogram. As a measure for the ellipticity of a contour, the decrease in FD₂ in the trained students reflects the more rounded ends of the average phonetogram (29). Comparing untrained and trained (singing) vocal groups in another study Sulter et al. (17) established a difference in FD₇. Future studies employing the shape analysis technique should confirm that a specific change in FDs reflects the nature and content of voice training.
Maximum phonation times

Maximum phonation times can be used to obtain valuable information about the voice function and expiratory system (19). Phonation time is determined both by the lung volume employed during phonation and the regulation of the generated airstream. Although differences in lung and, thus, phonation volumes exist among subjects, as this volume depends on gender, length and age, for a given phonation volume the phonation time is mainly determined by laryngeal resistance. Because in this study the same subjects were measured before (PRE) and after (POST) voice training, biasing influences of gender and length are not present. In addition, differences in maximum phonation times of vowels and voiced consonants are based on a modified regulation of both glottal resistance and expiratory control. In voiceless consonants, the maximum time is, apart from the phonation volume, dependent on expiratory control. Regression analysis in the described control cohort revealed no significant effect of age on maximum phonation times.

The significant decrease in phonation time of the voiced consonant /z/ suggests a decrease in glottal resistance. Without the presence of vocal fold irregularities, which was confirmed laryngostroboscopically, this decrement is related to more relaxed vocal folds, one of the goals of the trained exercises and methods. Large standard deviations are found PRE and POST for the /s/, expressing the large differences in expiratory control between subjects.

Previously the s/z ratio has been used to evaluate phonatory function, and possibly detect laryngeal pathology (20), or to follow the ageing process (35). Mueller (35) also refers to the sparsity of available normative data. The values established in this study for the /s/ and /z/ of the untrained students are only slightly shorter than those published in the study of Mueller (35). The s/z ratio, established in this study for untrained students, is slightly larger than the value reported by Mueller. All values (/s/, /z/ and s/z ratio) are close to those reported in the Eckel & Boone (20) study for a combined group of normal male and female subjects. In this study, a significant increase was found in the s/z ratio from 1.11 in the untrained situation to 1.45 after voice training. This increment is based on a decrease in phonation time of /z/ with an increase in /s/ value. In their article, Eckel & Boone (20) regard a larger than "normal" s/z ratio as a sign of laryngeal pathology. This study, however, shows that s/z ratios should be used with care and always related to absolute values of /s/ and /z/.

CONCLUSIONS

Specific voice training given to a group of untrained voice therapy students resulted in changes in both voice capacities and s/z ratio.
Analysis of differences in average phonetograms showed that trained students had increased dynamic ranges in the frequency range used in normal speech, and that soft voice capacities were greatly extended. Frequency ranges remained unchanged.

The s/z ratio increased significantly to a value heretofore associated with vocal pathology. This increase is presumably caused by reduced laryngeal resistance and improved expiratory control. S/z ratios should, therefore, be given with absolute values of /s/ or /z/ to distinguish between a normal subject having received vocal training leading to voluntarily reduced laryngeal resistance and improved expiratory control, and a patient with vocal pathology resulting in involuntarily reduced laryngeal resistance.

It is hypothesized that the increased dynamic capacities in the soft voice region are caused by the reduction of the threshold pressure, which relies on specific laryngeal features, such as vocal fold thickness and length, as well as on improved neuromyogenic control over breath support. The increased loud voice capacities could result from "resonance tuning" and flow phonation.

Voice training leads to a more efficient use of inspired air and optimizes both the sound generating capacities of vocal structures, as well as the resonance capacities of the vocal tract. Persons with limited natural capacities regarding sound production can therefore benefit from these functional changes, and dysfunctional voice use leading to pathology could be prevented. Prospective studies with different groups of vocal pathology should confirm the observations made on voice healthy subjects in this study.

References

Effects of Voice Training